BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 2: e26836
doi: 10.3897/biss.2.26836

OPEN ACCESS

Conference Abstract

The Death of VertNet Has Been Greatly Exaggerated But We Still Archive

Robert Guralnick[‡,§,|], John Wieczorek[¶], David Bloom[#]

‡ Vertnet, Florida, United States of America
§ University of Colorado, Boulder; University of Colorado Museum of Natural History, Boulder, United States of America
| Univ. of Florida, Gainesville, United States of America
¶ Museum of Vertebrate Zoology, University of California, Berkeley, United States of America
# VertNet, Berkeley, United States of America

Corresponding author: Robert Guralnick (rguralnick@flmnh.ufl.edu)

Received: 21 May 2018 | Published: 22 May 2018

Citation: Guralnick R, Wieczorek J, Bloom D (2018) The Death of VertNet Has Been Greatly Exaggerated But We Still Archive. Biodiversity Information Science and Standards 2: e26836. https://doi.org/10.3897/biss.2.26836

Abstract

VertNet (vertnet.org) is a collaborative project that makes biodiversity data free and available on the web. VertNet is also a tool designed to help people discover, improve, and publish biodiversity data. It is also the core of a collaboration between hundreds of biocollections that contribute biodiversity data and work together to improve it. VertNet has its genesis in the late 1990s and the very beginnings of vertebrate collections data sharing, and is nearing its 20th birthday. The small team that coordinates VertNet efforts long recognized the value of archival versions of VertNet data separate from individual published Darwin Core Archives. Here we describe why we produce what we call “snapshots” of the VertNet index. To understand the snapshots, it is important to also know how the VertNet indexing process works, which includes efforts at better flagging record types and special content of particular value to data consumers. We provide a brief explanation of the process we developed for creating these snapshots, focusing on how to assure their citation and licensing, and how to decide the scope of different snapshots. We also discuss the collaborative process of deciding infrastructure for archiving those snapshots, and our thinking about timing of new snapshots. In particular, we cover the use of Google BigQuery to produce snapshots and CyVerse as infrastructure for archival storage.

© Guralnick R et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords

VertNet, archiving, snapshots, vertebrate collections data, licensing, open data, data indexing

Presenting author

Robert Guralnick

Presented at

TDWG 2018